

(19)



Europäisches Patentamt  
European Patent Office  
Office européen des brevets

(11)



EP 0 986 046 A1

(12)

## EUROPEAN PATENT APPLICATION

(43) Date of publication:  
15.03.2000 Bulletin 2000/11

(51) Int Cl.7: G10H 1/00, G10H 3/12

(21) Application number: 99306932.7

(22) Date of filing: 31.08.1999

(84) Designated Contracting States:  
AT BE CH CY DE DK ES FI FR GB GR IE IT LI LU  
MC NL PT SE

Designated Extension States:  
AL LT LV MK RO SI

(30) Priority: 10.09.1998 US 150582

(71) Applicant: LUCENT TECHNOLOGIES INC.  
Murray Hill, New Jersey 07974-0636 (US)

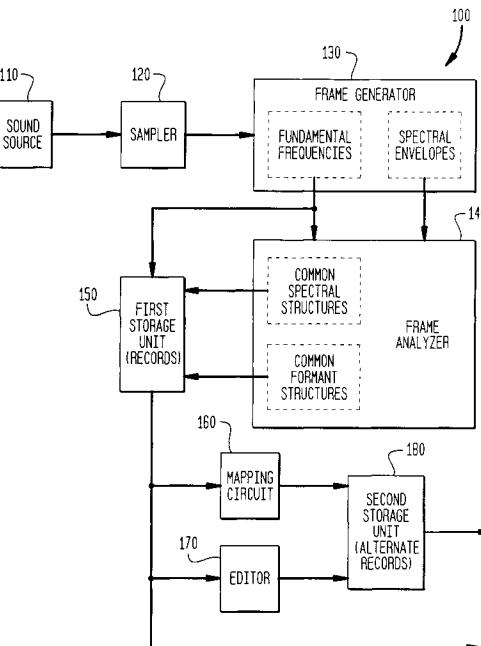
(72) Inventor: Curtin, Steven D.  
Freehold, New Jersey 07728 (US)

(74) Representative:  
Watts, Christopher Malcolm Kelway, Dr. et al  
Lucent Technologies (UK) Ltd,  
5 Mornington Road  
Woodford Green Essex, IG8 0TU (GB)

### (54) System and method for recording and synthesizing sound and infrastructure for distributing recordings for remote playback

(57) Systems and methods for recording and synthesizing sound in a resolution-independent manner and infrastructures for distributing resolution-independent recordings for remote playback. In one embodiment, one system includes: a frame generator (130) that extracts fundamental frequencies and spectral envelopes from the sound and creates frames therefrom and a frame analyzer (140) that identifies a selected one of common spectra structures and common formant structures in the frames and creates a record containing the fundamental frequencies and the selected one. One infrastructure includes: a radio station having a recording database associated therewith, a plurality of recordings contained within the recording database, each of the plurality of recordings including the selected one, a request receiver, coupled to the recording database, that receives remote requests for ones of the plurality of recordings and a transmitter, coupled to the recording database, that transmits the ones of the plurality of recordings in response to the requests.

FIG. 1



**Description****Technical Field Of The Invention**

**[0001]** The present invention is directed, in general, to sound recording and reproduction and, more specifically, to a resolution-independent system and method for making a reducing of a sound and later employing the record to synthesize the sound.

**Background Of The Invention**

**[0002]** Current recording of musical instrument performances are either based on sampling the analog signal of the instrument or recording the gestures that are input to a controller. This leads to playback situations where the performance may only be edited in a fixed time domain for the case of sampling, or a performance recording that may only realistically record keyboard and percussion in the case of Musical Instrument Digital Interface (MIDI). Additive and spectral synthesis technologies break down musical performances into discrete notes as opposed to a continuous performance. Generally, the ability to change the tempo or the key and to synchronize a performance with external events during playback are often difficult to accomplish at reasonable cost and without unacceptable distortions.

**[0003]** The sound waveforms produced may be characterized by many parameters, including frequency and amplitude. Using Fourier analysis, sound waveforms may be represented in a frequency domain as a spectral frame, consisting of spectral components. The spectral frame contains the waveform's lowest, or fundamental, frequency, along with its harmonics (spectral components which occur at multiples of the fundamental frequency). Spectral components from string instruments and from vowels in speech typically occur at close to whole number multiples of the fundamental frequency, while spectral components from percussion instruments often occur at non-integral multiples of the fundamental frequency.

**[0004]** Current sound recordings have been seen to be typically sample rate dependent or suffer from other recording and playback characteristics that make modifications to the record difficult to accomplish at acceptable costs or distortion levels. Additionally, these limitations make the offering of current sound recordings very limited in format for playback selection. Radio stations offer a selection of recordings that may be programmed for many days into the future with only an occasional specific request capability allowed, usually by telephone. Even in the case of selected requests, the recording is completely fixed in format with respect to tempo and key as well as its basic arrangement.

**[0005]** Therefore, what is needed in the art is a way to provide high quality, sample-rate independent recordings that may be selected in a random and expedient manner having the capability for specific listener modi-

fication or adaptation.

**Summary Of The Invention**

**[0006]** To address the above-discussed deficiencies of the prior art, the present invention provides systems and methods for recording and synthesizing sound in a resolution-independent manner and infrastructures for distributing resolution-independent recordings for remote playback. In one embodiment, one system includes: (1) a frame generator that extracts fundamental frequencies and spectral envelopes from the sound and creates frames therefrom and (2) a frame analyzer that identifies a selected one of common spectra structures and common formant structures in the frames and creates a record containing the fundamental frequencies and the selected one.

**[0007]** The present invention therefore introduces the broad concept of storing fundamental frequencies and selected structures in sound and creating a record containing those fundamental frequencies and selected structures to provide a basis for subsequent synthesis (tantamount to playback). The present invention preferably analyzes the sound as a continuous performance (irrespective of individual tones or notes).

**[0008]** In one embodiment of the present invention, the frames are discrete, allowing them to correspond to a discrete period of time. In an embodiment to be illustrated and described, common spectra or formant structures may be contained in a dictionary to compress the total size of the record.

**[0009]** In one embodiment of the present invention, a musical instrument generates the sound. Those skilled in the art are familiar with the formant content of certain musical instruments, such as string and wind instruments. Human voices likewise contain formants that may be captured and employed in later synthesis. The present invention can operate with any sound, however.

**[0010]** In one embodiment of the present invention, the frame generator samples the sound before extracting the fundamental frequencies therefrom. In the embodiment to be illustrated and described, sampling may occur at 1 ms intervals. However, those skilled in the art will understand that the present invention is not limited to a particular sampling frequency.

**[0011]** In one embodiment of the present invention, the system further includes a mapping circuit that applies a temporal quantization map to the record. Once the record is created, the present invention accommodates a wide range of conventional and later-developed sound manipulation techniques.

**[0012]** In one embodiment of the present invention, the system further includes an editor that modifies a selected one of a content and an order of the record. The editor allows still further manipulation of the sound once recorded.

**[0013]** In one embodiment of the present invention, the frame analyzer identifies the selected one of com-

mon spectra structures and common formant structures by Fourier analyzing the frames. Those skilled in the art are familiar with, in particular, fast Fourier transform techniques by which frequencies may be analyzed. The present invention is compatible with other conventional or later-developed spectrum analysis techniques, such as wavelets.

**[0014]** The present invention further provides infrastructures for distributing recordings for remote playback. One infrastructure includes: (1) a radio station having a recording database associated therewith, (2) a plurality of recordings contained within the recording database, each of the plurality of recordings including a selected one of common spectra structures and common formant structures corresponding thereto, (3) a request receiver, coupled to the recording database, that receives remote requests for ones of the plurality of recordings and (4) a transmitter, coupled to the recording database, that transmits the ones of the plurality of recordings in response to the requests.

**[0015]** The present invention therefore provides what amounts to "audio-on-demand" wherein formatted audio files are provided to remote "radios" to allow the remote "radios" to synthesize the audio in situ. Therefore, in one embodiment of the present invention, the infrastructure further includes a plurality of remote radios capable of receiving and digitally manipulating the ones of the plurality of recordings. The remote radios may comprise software that can be downloaded and executed on data processing and storage hardware to allow the ones of the plurality of recordings to be played. This infrastructure sharply contrasts with conventional analog AM or FM radio infrastructures in which remote radios simply demodulate and amplify received radio waves.

**[0016]** In one embodiment of the present invention, the transmitter broadcasts the ones of the plurality of recordings to receivers. Alternatively, the ones of the plurality of recordings may be addressed to individual remote "radios."

**[0017]** In one embodiment of the present invention, the ones of the pluralities of recordings are embodied in a plurality of bitstream files. The bitstream files contain data pertaining to the fundamental frequencies and the selected one as described above.

**[0018]** In one embodiment of the present invention, the recording database contains a record of the requests. This allows song popularity or advertisement dissemination to be tracked and accurate royalty payments to be calculated automatically.

**[0019]** The foregoing has outlined, rather broadly, preferred and alternative features of the present invention so that those skilled in the art may better understand the detailed description of the invention that follows. Additional features of the invention will be described hereinafter that form the subject of the claims of the invention.

### Brief Description Of The Drawings

**[0020]** For a more complete understanding of the present invention, reference is now made to the following descriptions taken in conjunction with the accompanying drawings, in which:

FIGURE 1 illustrates a block diagram of a resolution-independent system for recording a musical instrument constructed according to the principles of the present invention;

FIGURE 2 illustrates a flow diagram of a resolution-independent method of recording a musical instrument that may be carried out in the system of FIGURE 1;

FIGURE 3 illustrates a block diagram of a resolution-independent system for synthesizing a recorded musical instrument constructed according to the principles of the present invention;

FIGURE 4 illustrates a flow diagram of a resolution-independent method of synthesizing a recorded musical instrument that may be carried out in the system of FIGURE 3; and

FIGURE 5 illustrates a block diagram of a communications infrastructure capable of distributing resolution-independent recordings for remote playback.

### Detailed Description

**[0021]** Referring initially to FIGURE 1, illustrated is a block diagram 100 of a resolution-independent system for recording a musical instrument constructed according to the principles of the present invention. The resolution-independent system of block diagram 100 includes a sound source 110, a sampler 120, a frame generator 130, a frame analyzer 140, a first storage unit 150, a mapping circuit 160, an editor 170 and a second storage unit 180.

**[0022]** The present invention provides systems and methods for recording and synthesizing sound in a resolution-independent manner and infrastructures for distributing resolution-independent recordings for remote playback. In the present embodiment, a musical instrument may be used to generate the sound. Those skilled in the art are familiar with the formant content of certain musical instruments, such as string and wind instruments. Human voices also contain formants that may be captured and employed in later synthesis to be used in play-back of the recording. The present invention can operate with any sound, however.

**[0023]** The sampling, which may occur at 1 ms intervals, is accomplished by the sampler 120 before the fundamental frequencies are extracted by the frame gen-

erator 130. However, those skilled in the art will understand that the present invention is not limited to a particular sampling frequency or the use of a separate sampler as shown. The sampler 120 may be included as part of the frame generator 130. This embodiment allows the data comprising a sound source to be independent of the sampling rate or tempo desired. The frame generator 130 creates frames and extracts fundamental frequencies and spectral envelopes from the sound source 110 through the sampler 120. Then, the frame analyzer 130 identifies a selected one of common spectral structures and common formant structures in the frames and creates a record, containing this selected one and the appropriate fundamental frequencies, that is then stored in the first storage unit 150.

**[0024]** The present invention therefore introduces the broad concept of storing fundamental frequencies and selected structures in sound and creating a record containing those fundamental frequencies and selected structures to provide a basis for subsequent synthesis to be used in playback. The present invention analyzes the sound as a continuous performance, irrespective of individual tones or notes. The frames may be discrete, allowing them to correspond to a discrete period of time. The frame analyzer 140 identifies the selected one of common spectral structures and common formant structures by Fourier analyzing the frames. Those skilled in the art are familiar with Fast Fourier Transform (FFT) techniques as one technique by which frequencies may be analyzed. The present invention is compatible with other conventional or later-developed spectrum analysis techniques, such as wavelets, as well.

**[0025]** In this embodiment, common spectral or formant structures may be contained in a dictionary in order to compress the total size of the record. The identification and grouping of common spectral or formant structures allows them to be organized into a record structure that may be accessed in a way that is similar to words in a special dictionary to be used to reconstruct the particular sound composition. The dictionary may typically be a custom collection of common structures associated with the sound source being sampled, framed and analyzed. However, the dictionary may also contain common structures of a broader collection of sound sources that are recognized and tagged to correspond to the particular sound source being addressed. Further, there may be a collection of such dictionaries containing appropriate spectral and formant structures that are recognized and tagged.

**[0026]** The resolution-independent system of block diagram 100 further includes the mapping circuit 160 that applies a temporal quantization map to the record. Once the record is created, the present invention accommodates a wide range of conventional and later-developed sound manipulation techniques. The temporal quantization map allows the record to be synthesized with the ability to change tempo or key thereby providing different "feel factors". Additionally, it also provides the

ability to synchronize the playback performance with an external clock when other factors dictate.

**[0027]** The resolution-independent system of block diagram 100 still further includes the editor 170 that allows still further manipulation of the sound once recorded. The editor 170 may modify the content or the order of the record. The spectral structures or the formant structures may be modified, if desired, to provide effects not contained in the original sound source. The editor 170 may also be used to re-arrange the record sequence in time relative to the frequency contour. The resolution-independent system then stores these mapped or edited alternate records in the second storage unit 180.

**[0028]** Turning now to FIGURE 2, illustrated is a flow diagram 200 of a resolution-independent method of recording a musical instrument that may be carried out in the system of FIGURE 1. The flowchart 200 illustrates a method of recording sound which comprises extracting fundamental frequencies and spectral envelopes from the sound and creating frames from the fundamental frequencies and spectral envelopes. Then, one of common spectral structures and common formant structures in the frames are identified and selected to create a record containing both the fundamental frequencies and the ones selected.

**[0029]** The method begins in a start step 205 wherein the decision to create a record is made, and a sound source is selected in a step 210 that includes a musical instrument which generates the sound. The sound is sampled in a step 215 before extracting the fundamental frequencies from the sampled sound signal. Frames, that may be discrete, are then generated in a step 220 from which fundamental frequencies and spectral envelopes are then extracted. These frames are then analyzed in a step 225 using Fourier analysis, and then common spectral and formant structures are identified in a step 230. These common structures are then stored in a step 235 which creates a record of the common structures. A temporal quantization map may then be applied to the record as shown in a step 240, or the record may be modified selecting one of a content and an order of the record in a step 245 in order to edit the contents as required. The method ends in an end step 250 where the sound has been selected, sampled, framed, analyzed, recorded, mapped or modified in this embodiment.

**[0030]** Turning now to FIGURE 3, illustrated is a block diagram 300 of a resolution-independent system for synthesizing a recorded musical instrument constructed according to the principles of the present invention. The resolution-independent system of block diagram 300 includes a storage unit 305, a mapping circuit 310, an editor 315, a waveshaper 320, an output device 325 and a speaker 330. The storage unit 305 contains the records and dictionaries that have been created by sampling, framing and analyzing the sound source as discussed in FIGURE 1 and FIGURE 2 above.

**[0031]** The mapping circuit 310 applies a temporal quantization map to the record. As stated earlier, the present invention may accommodate a wide range of conventional and later-developed sound manipulation techniques allowing the record to be synthesized with a changed tempo or key and provide the ability to synchronize the playback performance with an external clock. Further manipulation of the recorded sound is provided with the editor 315 which may modify the content or the order of the record to provide effects not contained in the original sound source. The editor 315 may also be used to re-arrange the record sequence in time relative to the frequency contour as discussed.

**[0032]** The waveshaper 320 is coupled to the storage unit 305, the mapping circuit 310 and the editor 315. The waveshaper 320 takes the fundamental frequencies and applies a waveshaping transfer function to create a waveform from either the stored record, the mapped record or the edited record. The waveshaper 320 may also use some combination of these three in order to generate the waveform. The waveshaper 320 may also select from a number of waveshaping transfer functions that are stored in the waveshaper 320 to accommodate the waveshaping process. The waveshaper 320 is clocked externally, in this embodiment, allowing the waveform to be synchronized with external events. The waveform may then be converted into an output sound using the output device 325 and the speaker 330. The synthesis process represented here allows the originally recorded sound to be reproduced with appropriate fidelity or allows the originally recorded sound to be modified as deemed appropriate to the user. This does not preclude the use of other synthesis techniques such as FFT or direct sine.

**[0033]** Turning now to FIGURE 4, illustrated is a flow diagram 400 of a resolution-independent method of synthesizing a recorded musical instrument that may be carried out in the system of FIGURE 3. The method depicted in the flowchart 400 allows a temporal quantization map to be applied or a modification of the content or the order of a selected record to be accomplished in the creation and playback of sound recorded according to the present invention.

**[0034]** The method begins in a start step 405 wherein the decision to synthesize a record is made, and the record is selected in a step 410. The record is then processed using a temporal quantization map in a step 415. Then the record is modified through an edit function in a step 420. The selected, mapped and edited record is then waveshaped in a step 425 and the waveshaped record is then delivered for playback in a step 430. The method ends in an end step 435.

**[0035]** Turning now to FIGURE 5, illustrated is a block diagram 500 of a communications infrastructure capable of distributing resolution-independent recordings for remote playback. The block diagram 500 includes a wireless server 505 and an interactive music player 510 shown in FIGURE 5A, along with a random access play-

list 515, a download capability 520, a player 525 and a speaker 530 for sound reproduction shown in FIGURE 5B.

**[0036]** The present invention further provides infrastructures for distributing recordings for remote playback. One infrastructure includes the wireless server 505 depicted as a radio station having a recording database associated therewith, and a plurality of recordings contained within the recording database, where each of the plurality of recordings includes a selected one of common spectral structures and common formant structures corresponding to the records of each individual instrument or vocalist in a recording, which is re-synthesized and combined in real time during playback.

**[0037]** The interactive music player 510 generates requests to a corresponding request receiver that is coupled to the recording database, associated with the wireless server 505, which receives remote requests for various ones of the plurality of recordings. Additionally, a transmitter, coupled to the recording database, also associated with the wireless server 505, transmits the plurality of recordings in response to the requests. There is an additional monodirectional mode in which the receiver waits until the desired selection is transmitted before downloading and playing it.

**[0038]** The present invention therefore provides what amounts to audio-on-demand wherein formatted audio files are provided to remote radios or players, which act as "client" receivers, to allow these remote radios to synthesize the audio in situ. In this embodiment of the present invention, the infrastructure further includes a plurality of remote digital radios capable of receiving and digitally manipulating the plurality of recordings. This infrastructure sharply contrasts with conventional analog AM or FM radio infrastructures in which remote radios simply demodulate and amplify received radio waves. This infrastructure may function with any currently-proposed or later-developed digital transmitter and receiver standards. The program material for the recordings may include but is not limited to weather, news, stock quotes or other topical information.

**[0039]** The random access playlist 515 represents the pluralities of recordings, which are embodied in a plurality of bitstream files. The bitstream files contain data pertaining to the fundamental frequencies and the selected one as described above. The bitstream files, which may represent a collection of selections or the collection of offerings, may occur in a single serial loop. The user may select the ones of these that are downloaded and played. Alternately, they may occur in a collection of parallel loops allowing the user to perform the download 520 more rapidly. The transmitter, associated with the wireless server 505, may broadcast the ones of the plurality of recordings to all remote radios. Alternatively, the ones of the plurality of recordings may be addressed only to individual remote radios. The recording database may contain a record of the requests. This al-

lows song popularity or advertisement dissemination to be tracked and accurate royalty payments to be calculated automatically.

## Claims

1. A method of recording sound, for example sound produced by a musical instrument, comprising:
  - extracting fundamental frequencies and spectral envelopes from said sound;
  - creating frames from said fundamental frequencies and spectral envelopes;
  - identifying a selected one of common spectra structures and common formant structures in said frames; and
  - creating a record containing said fundamental frequencies and said selected one.
2. The method as recited in claim 1 wherein said frames are discrete.
3. The method as recited in claim 1 or claim 2 further comprising sampling said sound before extracting said fundamental frequencies therefrom.
4. The method as recited in any of the preceding claims further comprising applying a temporal quantization map to said record.
5. The method as recited in any of the preceding claims further comprising modifying a selected one of a content and an order of said record.
6. The method as recited in claim 8 wherein said identifying step includes Fourier analyzing said frames.
7. A system for recording sound, for example sound produced by a musical instrument, comprising means arranged to carry out a method as claimed in any of the preceding claims.
8. An infrastructure for distributing recordings for example recordings of sounds produced by musical instruments, for remote playback, comprising:
  - a radio station having a recording database associated therewith;
  - a plurality of recordings contained within said recording database, each of said plurality of recordings including fundamental frequencies and a selected one of common spectra structures and common formant structures corresponding thereto;
  - a request receiver, coupled to said recording database, that receives remote requests for ones of said plurality of recordings; and

a transmitter, coupled to said recording database, that transmits ones of said plurality of recordings in response to said requests.

- 5      9. The infrastructure as recited in claim 8 wherein said transmitter broadcasts said ones of said plurality of recordings to receivers.
- 10     10. The infrastructure as recited in claim 15 further comprising a plurality of remote radios capable of receiving and digitally manipulating said ones of said plurality of recordings.
- 15     11. The infrastructure as recited in any of claims 8 to 10 wherein said ones of said pluralities of recordings are embodied in a plurality of bitstream files.
- 20     12. The infrastructure as recited in any of claims 8 to 11 wherein said recording database contains a record of said requests.
- 25     13. A radio, comprising:
  - a receiver for receiving a recording including fundamental frequencies and a selected one of common spectra structures and common formant structures corresponding thereto;
  - a waveshaper, coupled to said receiver, for applying a waveshaping transfer function based on said selected one to said fundamental frequencies to create a waveform; and
  - a speaker, coupled to said waveshaper, for converting said waveform into an output sound.
- 30     14. The radio as recited in claim 13 further comprising a mapping circuit, coupled to said receiver, that applies a temporal quantization map to said selected one.
- 35     15. The radio as recited in claim 13 or claim 14 wherein said waveshaping transfer function is selected from a plurality of waveshaping transfer functions stored in said waveshaper..
- 40     16. The radio as recited in any of claims 13 to 15 wherein said waveshaper is clocked externally.
- 45     17. The radio as recited in any of claims 13 to 16 further comprising an editor, coupled to said receiver, for modifying a content of said recording.

FIG. 1

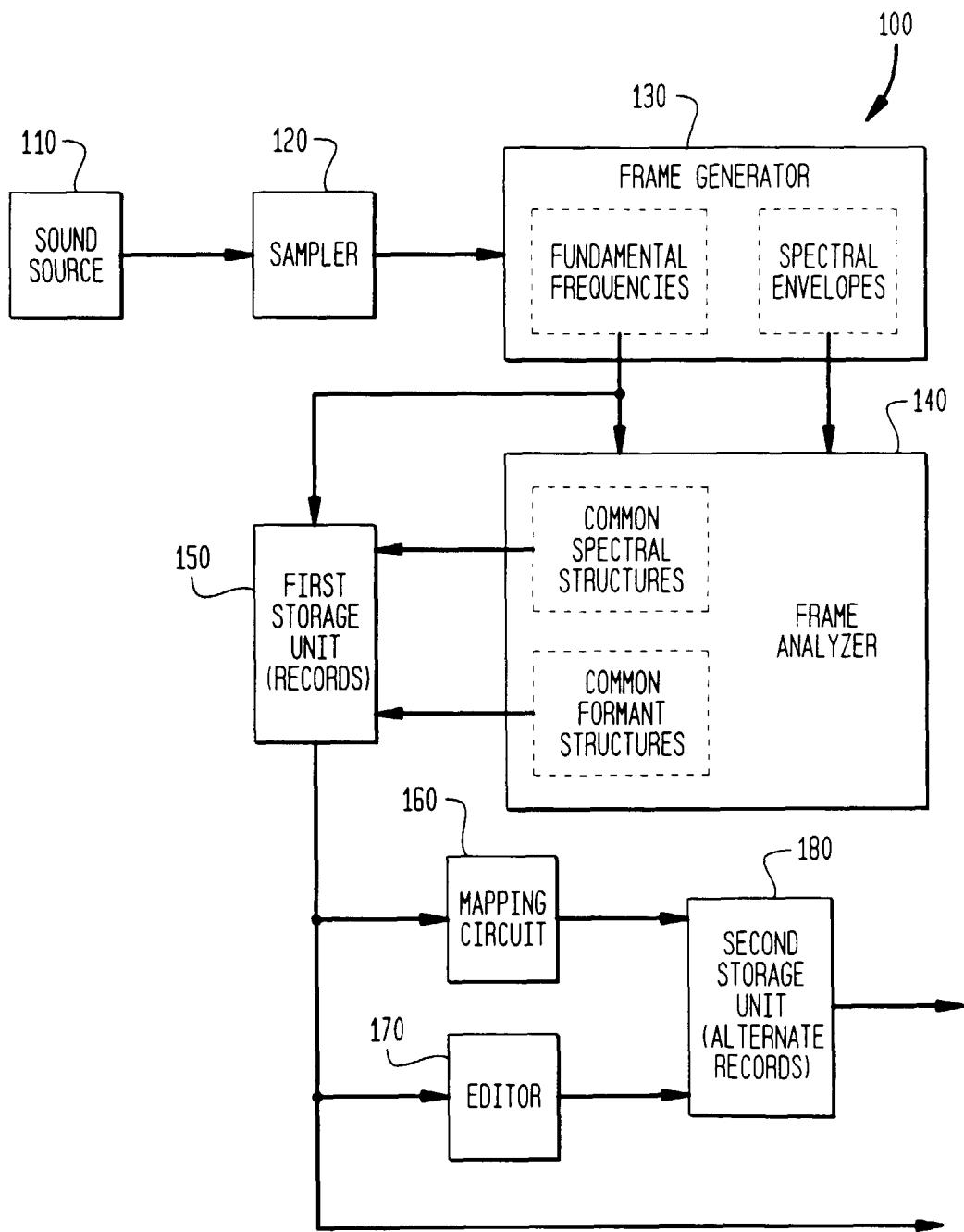


FIG. 2

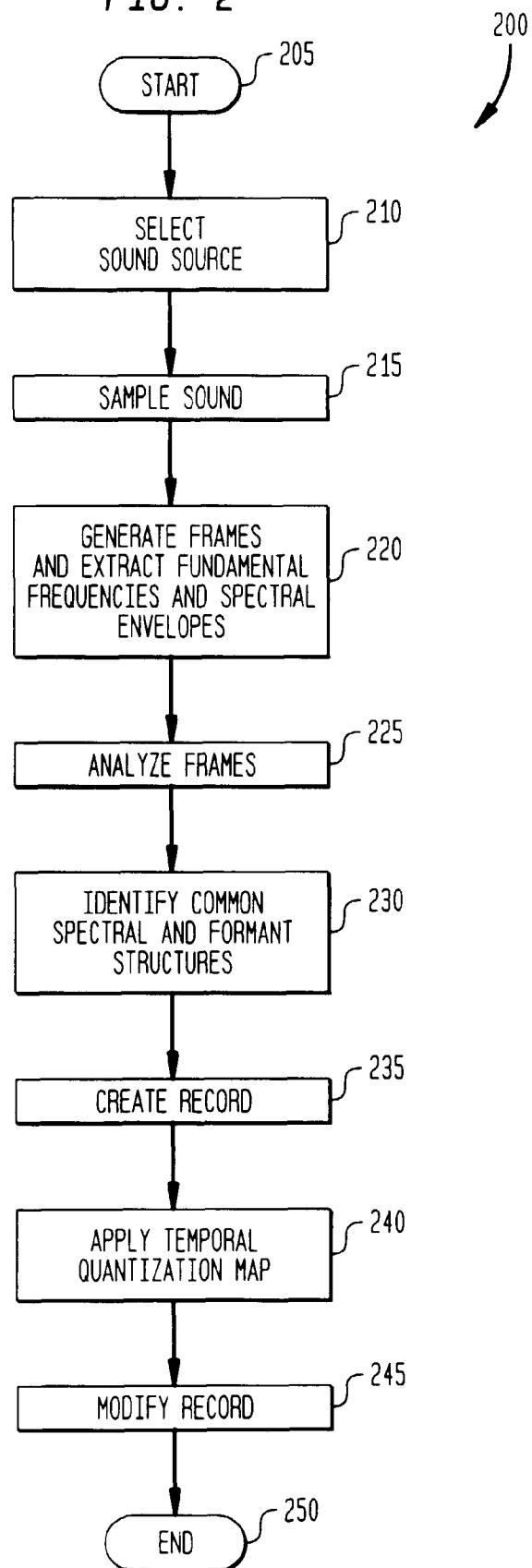


FIG. 3

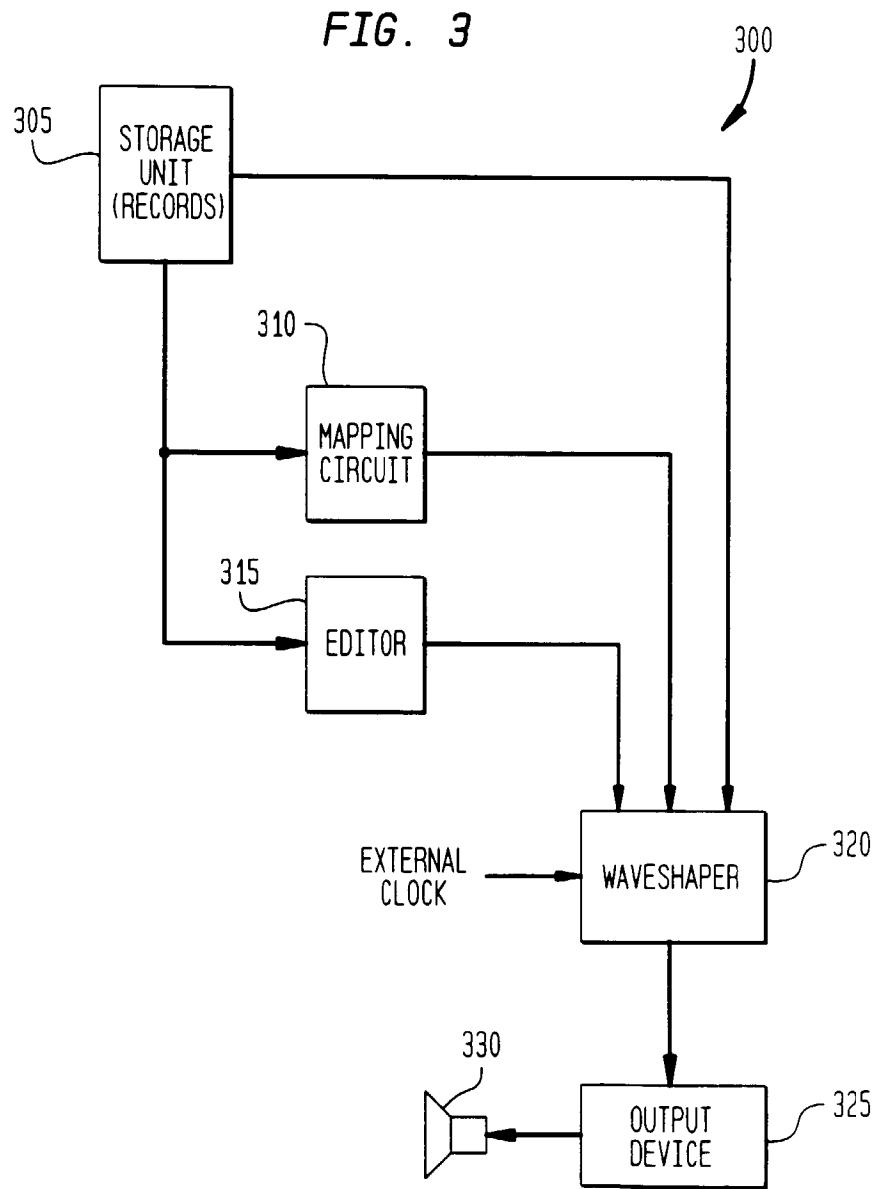


FIG. 4

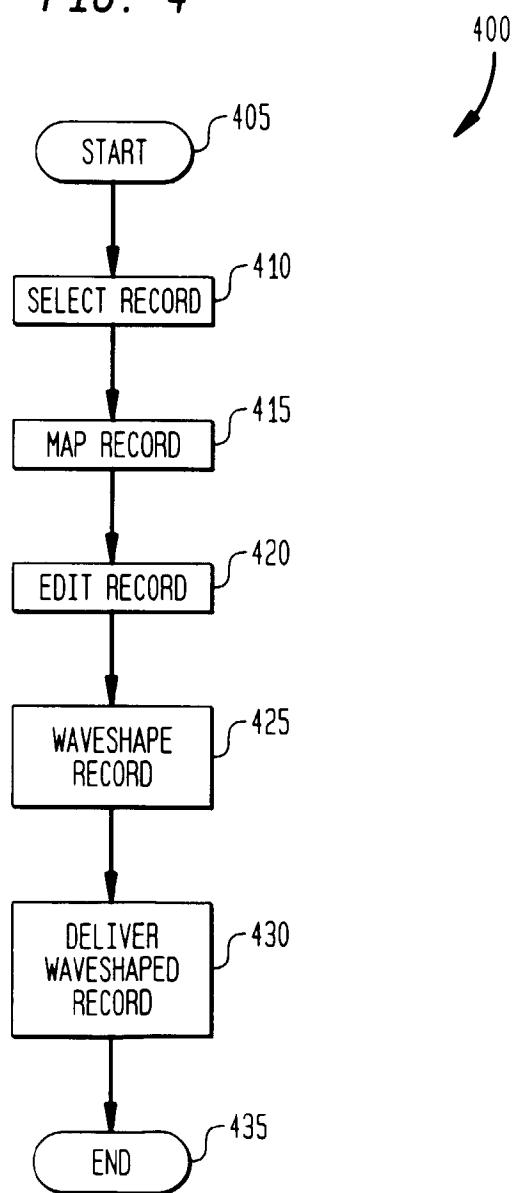


FIG. 5A

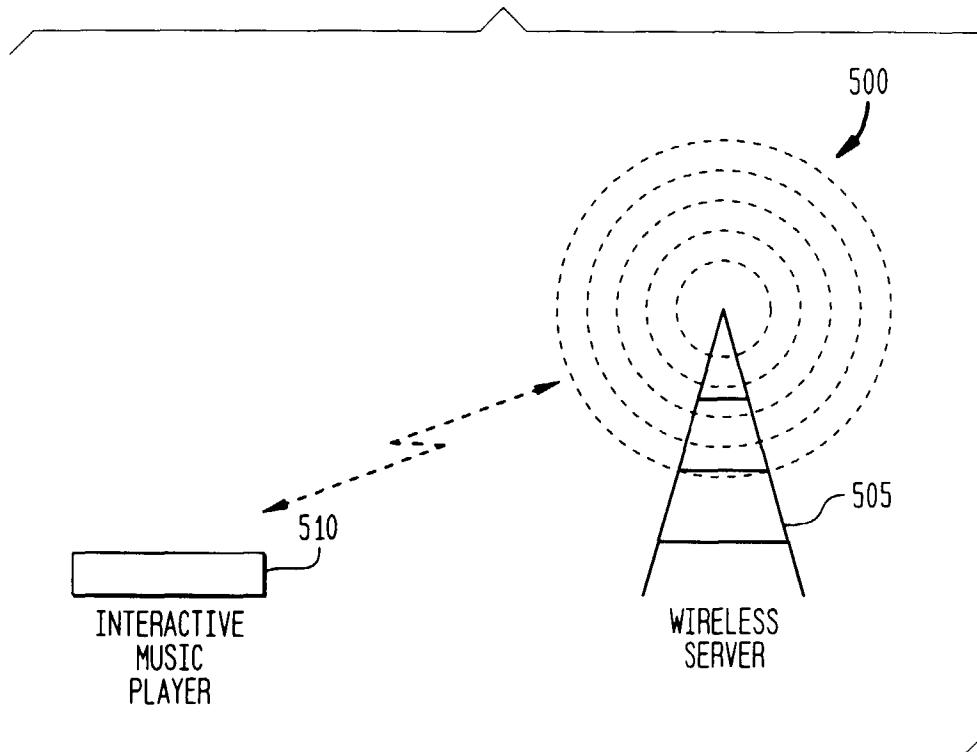
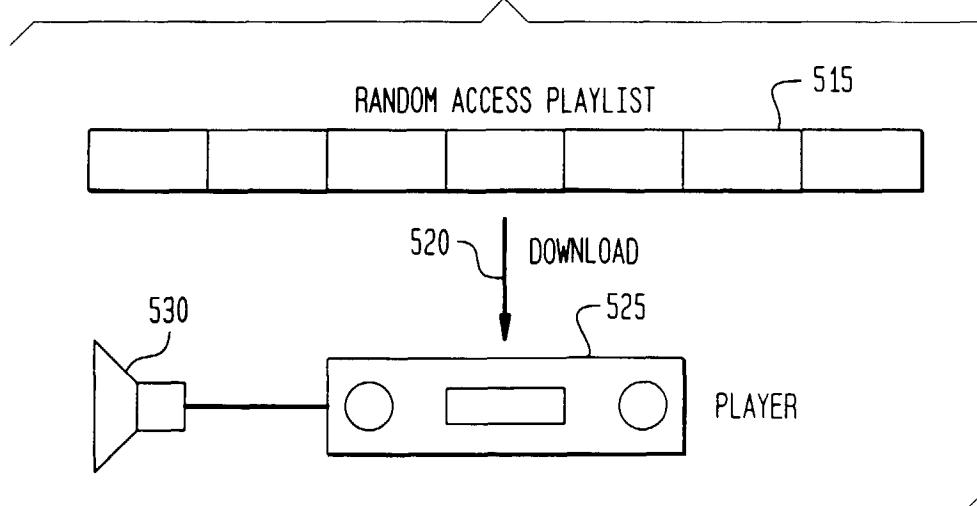


FIG. 5B





European Patent  
Office

## EUROPEAN SEARCH REPORT

Application Number  
EP 99 30 6932

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int.Cl.7)
X	US 5 536 902 A (WOLD ERLING ET AL) 16 July 1996 (1996-07-16) * column 3, line 7 - line 46 * * column 9, line 58 - column 10, line 67; figure 2 *	1-6	G10H1/00 G10H3/12
P,X	US 5 886 276 A (LEVINE SCOTT N ET AL) 23 March 1999 (1999-03-23) * column 4, line 29 - column 6, line 67 * * column 10, line 52 - column 11, line 5 * * column 13, line 22 - column 14, line 25; figures 1,5 *	1-17	
A	EP 0 686 923 A (MOTOROLA INC) 13 December 1995 (1995-12-13) * page 3, line 8 - line 34 *	1,6,8	
A	US 5 248 845 A (MASSIE DANA C ET AL) 28 September 1993 (1993-09-28) * column 4, line 3 - line 21 * * column 5, line 11 - column 6, line 10 * * column 16, line 46 - line 68; figure 2 *	1-7	
A	US 5 735 744 A (OKAMOTO TAKEYA) 7 April 1998 (1998-04-07) * column 9, line 4 - line 54 * * column 11, line 43 - line 59; figures 1,2,10,11 *	8-17	G10H
The present search report has been drawn up for all claims			
Place of search	Date of completion of the search	Examiner	
THE HAGUE	29 November 1999	Pulluard, R	
CATEGORY OF CITED DOCUMENTS			
X : particularly relevant if taken alone	T : theory or principle underlying the invention		
Y : particularly relevant if combined with another document of the same category	E : earlier patent document, but published on, or after the filing date		
A : technological background	D : document cited in the application		
O : non-written disclosure	L : document cited for other reasons		
P : intermediate document	& : member of the same patent family, corresponding document		

**ANNEX TO THE EUROPEAN SEARCH REPORT  
ON EUROPEAN PATENT APPLICATION NO.**

EP 99 30 6932

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report. The members are as contained in the European Patent Office EDP file on. The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

29-11-1999

Patent document cited in search report		Publication date		Patent family member(s)	Publication date
US 5536902	A	16-07-1996		JP 2906970 B JP 7325583 A	21-06-1999 12-12-1995
US 5886276	A	23-03-1999		NONE	
EP 0686923	A	13-12-1995		US 5473557 A JP 8006924 A	05-12-1995 12-01-1996
US 5248845	A	28-09-1993		AU 3918293 A WO 9319455 A US 5698807 A	21-10-1993 30-09-1993 16-12-1997
US 5735744	A	07-04-1998		JP 2532198 B JP 6334780 A US 5775995 A US 5489103 A	11-09-1996 02-12-1994 07-07-1998 06-02-1996